Cognition 132 (2014) 151–163

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

Object labeling influences infant phonetic learning and generalization

H. Henny Yeung *, Thierry Nazzi

CNRS (Laboratoire Psychologie de la Perception, UMR 8242) & Université Paris Descartes, Paris Sorbonne Cité, France

ARTICLE INFO

Article history:
Received 6 August 2013
Revised 31 March 2014
Accepted 1 April 2014
Available online 5 May 2014

Keywords:
Infant
Speech perception
Word learning
Prosody
Lexical stress

ABSTRACT

Different kinds of speech sounds are used to signify possible word forms in every language. For example, lexical stress is used in Spanish (/'be.be/, '*he/she drinks*' versus /be.'be/, '*baby*'), but not in French (/'be.be/ and /be.'be/ both mean '*baby*'). Infants learn many such native language phonetic contrasts in their first year of life, likely using a number of cues from parental speech input. One such cue could be parents' object labeling, which can explicitly highlight relevant contrasts. Here we ask whether phonetic learning from object labeling is abstract—that is, if learning can generalize to new phonetic contexts. We investigate this issue in the prosodic domain, as the abstraction of prosodic cues (like lexical stress) has been shown to be particularly difficult. One group of 10-month-old French-learners was given consistent word labels that contrasted on lexical stress (e.g., Object A was labeled /'ma.bu/, and Object B was labeled /ma.'bu/). Another group of 10-month-olds was given inconsistent word labels (i.e., mixed pairings), and stress discrimination in both groups was measured in a test phase with words made up of new syllables. Infants trained with consistently contrastive labels showed an earlier effect of discrimination compared to infants trained with inconsistent labels. Results indicate that phonetic learning from object labeling can indeed generalize, and suggest one way infants may learn the sound properties of their native language(s).

1. Introduction

Experience with one's native language(s) alters speech perception from early in infancy (Jusczyk, 2000). This developmental phenomenon is well exemplified by *phonetic attunement*, which describes a decline in infants' perception of certain non-native phonetic contrasts and improvement in the perception of many other phonetic contrasts—particularly native ones—a process that begins already in the first year of life (e.g., Best, McRoberts, & Goodell, 2001; Eilers, Wilson, & Moore, 1977; Narayan, Werker, & Beddor, 2010; Rivera-Gaxiola, Klarman, Garcia-Sierra, & Kuhl, 2005; Werker & Tees, 1984; Yeung, Chen, & Werker, 2013). Many of the phonetic patterns to which infants attune are determined by the lexical structure of the native language, but it is important to note that attunement begins *before* infants have acquired lexicons of any substantial size. Recent work has thus focused on learning mechanisms that could drive phonetic attunement without taking lexical knowledge into account.

One such mechanism suggests that infants establish phonetic categories based on regions of acoustic space in which the most frequently heard speech tokens occur (Anderson, Morgan, & White, 2003; Jusczyk, 1993; Kuhl et al., 2008). For example, in a process termed *distributional learning*, infants are supposed to identify native language phonetic categories by tracking the frequency distributions

* Corresponding author. Address: Laboratoire Psychologie de la Perception, 45 rue des Saints-Pères, 75006 Paris, France. Tel.: +33 0 1 42 86 21 96.

E-mail address: henny.yeung@parisdescartes.fr (H.H. Yeung).

http://dx.doi.org/10.1016/j.cognition.2014.04.001
0010-0277/

of phonetic tokens in acoustic space (Salminen, Tiitinen, & May, 2009; Toscano & Mcmurray, 2010; Vallabha, McClelland, Pons, Werker, & Amano, 2007; Werker et al., 2007). A number of studies have shown that infants' perceptual sensitivities in the speech domain are indeed affected by the distributional characteristics of speech input (Cristia, 2011; Cristia, McGuire, Seidl, & Francis, 2011; Maye, Weiss, & Aslin, 2008; Maye, Werker, & Gerken, 2002; Yoshida, Pons, Maye, & Werker, 2010). Nevertheless, recent work has challenged the notion that distributional learning alone can explain phonetic attunement (see Werker, Yeung, & Yoshida, 2012 for review). For example, it is unclear whether distributional information can support the learning of acoustically difficult phonetic contrasts (Cristia et al., 2011; Narayan et al., 2010; Sato, Kato, & Mazuka, 2012), or explain learning when phonetic distributions show substantial overlap (Feldman, Myers, White, Griffiths, & Morgan, 2013; Swingley, 2009). Indeed, other computational models have suggested that additional learning mechanisms beyond simple token-counting—such as an unsupervised competitive learning process that selects among possible statistical distributions—are needed to successfully learn some phonetic distinctions (McMurray, Aslin, & Toscano, 2009; Toscano & Mcmurray, 2010).

Infants may also use cues from the rich environment in which speech is embedded to either constrain or supplement statistical learning. For example, phonetic categories are better learned in "social" situations involving live, interactive experimenters compared to video presentations of these same interactions (Kuhl, Tsao, & Liu, 2003). Visual information from talking faces can also affect infants' distributional learning abilities (Teinonen, Aslin, Alku, & Csibra, 2008), and visual cues from the *lexical context* in which speech is heard (e.g., hearing "ball" in the context of a ball, and "doll" in the context of a doll) may also have an influence. For example, previous reports in both adults (Hayes-Harb, 2007) and infants (Yeung & Werker, 2009) show that repeated presentations of such lexical contexts can increase the perceptual distance between the target phonetic categories. Lexical contexts may also be auditory: Repeated presentations of [i] in the context *see* and [I] in the context *this* can increase the perceptual distance between those vowels (Feldman et al., 2013; Swingley, 2009; Thiessen, 2007).

Phonetic learning from visual and auditory contexts such as these illustrates a more domain-general associative process called *acquired distinctiveness*. Here, the perceptual distance between two stimuli (e.g. two phonetic tokens) is increased by the unique presentation of these stimuli in different contexts (Kluender, Lotto, Holt, & Bloedel, 1998; Lawrence, 1949). The current study focuses on learning from acquired distinctiveness, and asks whether a newly learned phonetic distinction exemplified by two contrasting object labels is abstract—that is, if a newly learned phonetic contrast can be *generalized* to a different context.

1.1. Generalization in phonetic learning

Generalization is highly relevant to models of speech processing and representation, as well as to theories of phonological learning (see Cristia, Seidl, & Francis, 2011 for review). Research on phonological processing has shown, for example, that infants not only learn constraints on the location of sounds in certain positions in the syllable (phonotactic regularities), but also generalize this phonotactic knowledge to new contexts from 4 to 16.5 months of age (Chambers, Onishi, & Fisher, 2003, 2011; Cristia & Seidl, 2008; Seidl, Cristia, Bernard, & Onishi, 2009).

The generalization of *phonetic category* knowledge is also the topic of several studies. In adults, for example, lexical training can bias the perception of phonetic information towards one category versus another (i.e., accent adaptation), and this training effect generalizes across both items (Norris, McQueen, & Cutler, 2003) and speakers (Kraljic & Samuel, 2006). In a related set of studies, toddlers from 18 to 30 months of age have also been shown to adapt to accentual differences in the pronunciation of vowels, and generalize these accentual patterns when identifying familiar words (White & Aslin, 2011), or when learning novel words (Schmale, Cristia, & Seidl, 2012). Although this kind of phonological generalization is robust, occurring even after brief amounts of exposure in the laboratory (see also Schmale, Hollich, & Seidl, 2011; van Heugten & Johnson, 2014), the ability to generalize phonetic category information across accents in *younger* infants—still in their first year of life—is less clear, as learners this age are still building stable phonological representations that must be robust enough to either be generalized (in the case of accentual differences) or not be generalized (in the case of different vowels), depending on the appropriate contexts (Best, Tyler, Gooding, Orlando, & Quann, 2009; Mulak, Best, Tyler, Kitamura, & Irwin, 2013; Schmale & Seidl, 2009).

To our knowledge, only one previous report has examined the issue of generalization in young infants' learning of phonetic categories (Maye et al., 2008). This study showed that English-learning 6- to 8-month-olds could be trained with a statistical distribution of phonetic tokens that indicated two non-native phonetic categories (a prevoiced coronal stop [da] versus a voiceless unaspirated coronal stop [ta]), and then generalize this learning to a different phonetic class (the same voicing distinction, but with velar stops: [ga] versus [ka]). This is an intriguing example of infants generalizing a phonological feature (voicing) to new phonetic contexts (across coronal and velar phonetic classes).

The current experiment expands on Maye et al.'s (2008) classic generalization study in two important ways. The first extension was to investigate whether phonetic generalization can be driven by another phonetic learning mechanism. Here we focused on acquired distinctiveness by creating two learning situations where information from statistical distributions was similar, but where information about lexical structure (i.e., the pairings of objects and labels) was not. Infants were then tested on discrimination of the target contrast in a novel phonetic context, and differences as a function of training would suggest that labeling could supplement both the learning and generalization of a phonetic distinction.

The second extension was to examine a new phonetic contrast: French-learning infants' perception of lexical stress. On the one hand, this stress contrast is similar to